

104

TP183
Z32

二进前向人工神经网络 ——理论与应用

张军英 许 进 著



A0956627

西安电子科技大学出版社

2001

图书在版编目(CIP)数据

二进前向人工神经网络:理论与应用/张军英,许进著.

—西安:西安电子科技大学出版社,2001.5

高等学校教材

ISBN 7-5606-1006-4

I. 二… I. ①张… ②许… III. 人工神经网络-高等学校-教材 IV. TP183

中国版本图书馆 CIP 数据核字(2001)第 08933 号

责任编辑 梁家新 李纪澄

出版发行 西安电子科技大学出版社(西安市太白南路2号)

电 话 (029)8227828 邮 编 710071

http://www.xduph.com E-mail: xdupfb@pub.xaonline.com

经 销 新华书店

印 刷 西安电子科技大学印刷厂

版 次 2001年5月第1版 2001年5月第1次印刷

开 本 787毫米×1092毫米 1/16 印张 12.875

字 数 286千字

印 数 0 001~2 000册

定 价 21.00元

ISBN 7-5606-1006-4/TP·0490

*** 如有印装问题可调换 ***

本书封面贴有西安电子科技大学出版社的激光防伪标志,无标志者不得销售。

序

在科学技术高度发展的今天，人类对于世界的认识已经达到了令人惊叹的程度，无论是宏观远及亿万光年之大的宇宙，还是微观深入到层子、夸克的神奇世界。人脑进行信息加工和处理所凭借的基本结构是由大约 10^{11} 个神经元及其突触构成的生物神经网络。生物神经网络在人脑信息处理过程中的重要地位和它巨大的处理能力，使得人们对以模拟生物神经网络的人工神经网络给予了极大的关注，特别是超大规模集成电路技术的迅猛发展，为人工神经网络的技术实现提供了强大的手段。因此，虽然目前人工神经网络还只是对高级生物神经网络的极其简化的模型表示，却具有非常重要的意义。

以冯·诺依曼体系结构为基础的所有现代计算机，在数字运算方面所表现出的能力和效率已经达到了人脑望尘莫及的程度，但它在处理模式信息方面的能力甚至还远不如一个婴儿，而模仿生物神经网络的人工神经网络，是一种具有大规模并行处理、分布式信息存储和高度容错能力的具有自学习、自组织、自适应的非线性网络，它与冯·诺依曼结构截然不同的模拟生物神经网络的结构和运行方式，无疑给探求以高效率处理模式信息的新一代计算方式和新一代计算机的科学家们看到了希望的曙光。

前向人工神经网络作为人工神经网络的典型模型之一，在模式信息处理和模式分类方面研究得最早，应用得最为广泛，对促进人工神经网络在 20 世纪 80 年代的复苏起到了重要作用。然而到目前为止，对于最为典型的前向人工神经网络——二进前向人工神经网络的研究，仍存在许多有待解决的基本理论问题，这些问题的研究和解决将对前向人工神经网络的研究和应用起到重要的促进作用。

张军英、许进是两位年轻的博士，近年来致力于二进前向人工神经网络的研究，这本著作是他们在认真阅读有关文献资料基础上，经过多年艰苦细致的研究工作写作而成的，反映了他们在这些方面的研究成果。可以看出，作者收集、整理和阅读了大量文献资料，在二进前向人工神经网络方面进行了许多有

意义的探索性工作，取得了可喜的成绩，使得这本书有以下几个方面的特点：

(1) 内容较丰富，有一定的前沿性。书中仅用很少篇幅描述神经网络的一般知识，而用大量篇幅对二进前向人工神经网络进行较为全面系统的讨论，包括二进前向人工神经网络的基本理论、基本性质、网络的分类机理、网络的结构优化、网络的设计和训练等，其中大部分内容是作者近年来的研究成果。

(2) 分析较严谨。书中对问题的讨论深入细致，并附有严格的理论推导，同时有较详细的评注和说明，有助于读者对问题的深入分析、理解和认识。

(3) 逻辑性较强。书中以二进前向人工神经网络的分类能力为主线，基于最稳健的网络才可能达到网络资源的最有效利用及网络规模最小的思想，将网络分类能力和网络稳健性能在统一的框架下展开研究，使得书中的内容具有较好的逻辑性。

这本书的上述特点使得它适合于人工神经网络理论和应用研究方面的科技人员和研究生阅读，它的出版无疑会对人工神经网络的研究起到一定的促进作用。希望作者和读者不断努力，为我国人工神经网络学科的发展作出新的贡献。

保 铮 谨识

2001 年 4 月于西安电子科技大学

自 20 世纪 40 年代信息科学的开创时期诞生到现在,人工神经网络的研究已经走过了半个多世纪的曲折历程,吸引了许多领域科学家的重视和研究,有关的杂志、会议不断涌现,成为现代脑神经科学、数理科学以及信息科学等综合研究领域的共同的科学前沿之一。

前向网络是典型的人工神经网络模型之一,也是在模式识别和分类方面发展得最早、研究人员最多、应用最为广泛的一类人工神经网络模型,在工程上有着重要的应用和前景。二进前向网络是一种典型的离散型前向网络,而现在的计算机中存在着大量逻辑电路,因此,单就其可以以比逻辑电路高得多的效率实现布尔函数这一点看,二进前向网络的研究就具有非常重要的理论和实际意义。此外,就前向网络而言,如果将其输入至第一隐层视为是对输入空间的矢量量化过程,则其第一隐层输出至网络输出所完成的实际上就是一个二进前向网络的功能。再者,二进前向网络在逻辑知识推理、模式分类与识别、非线性预测、故障诊断、自动控制等方面都有重要的理论和应用研究意义。

作者张军英 1992 年申请并承担了国家(青年)自然科学基金项目“基于人工神经网络的说话人识别研究”,致力于将人工神经网络技术应用于说话人识别的研究。从该项研究过程中作者深深体会到,人工神经网络无论用于何方面的研究,开始人们认为它可以无所不用,但随着应用的不断深入,其理论研究不够深入所导致的应用上缺乏理论根据和指导的弱点也不断暴露出来了,极大地阻碍了人工神经网络的有效应用。1995 年张军英有幸攻读了中国科学院院士保铮教授的在职博士,并从那时起,在保铮教授的悉心指导和支持下,开始了对二进前向网络的比较深入的研究,本书即是在此基础及其后续工作基础上产生的。

本书是关于二进前向网络方面的第一部学术专著,其中的大部分内容为作者有关二进前向网络方面的研究成果。全书以二进前向网络的分类能力为主线,基于最稳健的网络才可能达到网络规模最小、网络资源利用率最高的思想,将网络的分类能力和网络的稳健性能统一考虑。全书共分八章,系统地研究了二进前向网络的基本性质和基本理论,诸如感知器及其学习;稳健感知器及其学习;实现 Boole 函数的稳健和最稳健二进前向网络的设计;稳健二进前向网络的基本性质,包括网络的连接特性、辅阈值的区域特性和辅阈值的反向贡献特性等;稳健二进前向网络用于逻辑知识的获取、存储、推理和解释;网络隐层稳健神经元对应的稳健分类超平面系的基本模式(平行模式、

包围模式)及其(稳健)线性可分性,一般模式(不相交模式、可相交模式)及其(稳健)线性可分性;通过合并隐层神经元以达到减少网络隐节点数目、降低网络规模的分类超平面系的退化和稳健退化的条件;二进前向网络的训练和学习算法及其收敛性等。

本书由张军英、许进共同撰写。其中,第1章的1.4节、第2章的2.1、2.2、2.3节以及附录A中的A1、A2、A3、A4节均由许进撰写,其余章节由张军英撰写。

我们在前向网络方面所开展的研究工作得到了两项国家自然科学基金项目的资助,并得到了北京大学视觉与听觉信息处理国家重点实验室开放基金至今连续六年的资助,在此表示衷心的感谢。我们特别感谢我们的恩师——中国科学院院士保铮教授几年来对我们的支持、关怀、指导和帮助,同保老师的每次讨论都使我们受益匪浅,保老师严谨科学的治学态度一直影响着我們。感谢西安电子科技大学雷达信号处理国防重点实验室的吴顺君所长及全所各位老师对我们的大力支持;感谢该重点实验室冯大政教授、水鹏朗副教授和作者所在课题组讨论班的所有同志,与他们的讨论对于完善我们的工作起了重要作用。

另外,我们衷心感谢曾任国际神经网络学会主席、世界著名的日本RIKEN研究所的Amari教授,在国际会议期间与他的交流以及在国际会议之后他专门写给作者张军英的memo,对我们的工作有非常大的启发;感谢新西兰Otago大学信息科学系的Kasabov教授,在国际会议期间与他的交流以及在此期间,把他们研制的有关软件和书稿近作拷贝给作者张军英,这对我们的工作是一个很好的参考。

感谢复旦大学数学系陈天平教授、西南交通大学靳番教授和复旦大学电子工程系张立明教授,他们在国内和国际会议期间对我们的科研工作提出了很好的建议。衷心感谢厦门大学物理系的徐慎初教授,在国内、国际会议期间多次牺牲休息时间与作者张军英进行了长时间的专题讨论。在本书的写作和完成过程中,西安电子科技大学出版社李荣才总编辑、梁家新副社长对出版本书给予了极大的帮助,在此一并表示感谢。

由于作者学识和水平有限,书中不妥之处在所难免,恳请读者给予批评指正。

作 者

2000年12月20日

于西安电子科技大学

Preface

The study of artificial neural networks has ever been evolving for about half a century and it has attracted a large number of researchers in many different areas. Now artificial neural networks have become a frontier of encephaloneural science, mathematics and information science.

Feedforward neural network is one of the most important type of artificial neural networks. It find wide application in engineering, such as pattern classification and pattern recognition. Binary feedforward neural network, its inputs and outputs being all binary numbers, is very perspective in engineering, since it could even realize Boolean functions much more efficiently than a logic circuit does, while the computer today consists of a huge number of logic circuits. For a feedforward neural network, if the first hidden layer of the network is thought of as a nonlinear mapping or a vector quantilization from its continuous input space of the problem to a binary valued space, the remaining part of the network from its second hidden layer to its output layer could be thought of as a binary feedforward neural network. On the other hand, binary feedforward neural network is very applicable and efficient in inference of logic rules, pattern classification and recognition, nonlinear prediction, fault diagnoses, automatic control and so forth.

The first author of the present book began her research in artificial neural networks by taking responsibility for a national natural science fund on mentioned item in 1992. She was engaged in application of artificial neural network to speaker recognition. From her own experience acquired in the process of deep study of artificial neural networks in respect to their applications, the author recognizes the following truth. In studying application of the artificial neural network in whichever area people at first thought that it could be applied everywhere. With the more applications of it the weak points in its study become obvious, that is, the study lacks of systematic and theoretic analysis of the behavior of the network. These shortcomings obstruct its further application to a great extent. That is why in 1995 the first author began to specialize in artificial neural networks as a Ph. D. candidate and has obtained later some results on them. This book is formulated in abovementioned basis.

This is the first monograph which gives systematic presentation of binary feedforward neural networks. Most of its contents is the research results performed by the authors. The classification capability and robust performance of the networks are considered as a whole on the basis of the fact that only when the network is the most robust and in its smallest scalability-

ty, it can be made the most efficient in its resource utilization. The book is divided into eight chapters. Systematic study is given of basic properties and fundamental theories of binary feedforward neural networks, such as perceptron and its learning; robust perceptron and its learning; robust and the most robust design of binary feedforward neural networks to knowledge extraction, storage, inference and interpretation; formats of a system of classification hyperplanes in input space and their separability, degeneration and robust degeneration of a system of classification hyperplanes to decrease the number of hidden neurons in the networks; the training and learning algorithms and their convergence of binary feedforward neural networks.

The book is written by Zhang Junying and Xu Jin. Except sections 1.4, 2.1, 2.2, 2.3 and appendix which are written by Xu, the rest chapters and sections are written by Zhang. The authors hope that the book will be proved useful to those students and practicing professionals who are interested not only in understanding the underlying theory of feedforward neural networks but also in pursuing research in this area.

The authors would firstly like to acknowledge the National Natural Science Fund Committee for the support with two projects on artificial neural networks. Sincere thanks are offered to the National Key Laboratory on Machine Perceptron of Beijing University for its fund offered for six years so far. A special gratitude is expressed to Prof. Bao Zheng, our adviser, academician in China Science Research Institute, for many enjoyable and productive consultations. The authors are also grateful to Key Laboratory of Radar Signal Processing for its supports. Thanks are due to all of the colleagues in forum for their great help in improvement of the research work. Thanks go to Prof. Ameri, who was ever chairman of the International Neural Network Society, for that the discussion with him during an international conference and the memo he wrote and emailed to the first author after the conference enlightened the author on her work very much. Thanks must be given to prof. Kasabov in Information Science Department of Otago University, New Zealand. During an international conference he copied the software on feedforward neural networks they programmed and the materials of the work he just completed and gave them to the author. Both of the gifts are very useful references to the authors' research work. Thanks must also be extended to Prof. Chen Tianping of Mathematics Department of Fudan University, Prof. Jin Fan in Southwest Jiaotong University and Prof. Zhang Liming in Electronics Engineering Department of Fudan University, all of them gave the author many suggestions and encouragement during national and international conferences. Deep gratitude goes to Prof. Xu Shenchu of Physics Department of Xiamen University, who had several discussions with the first author during conferences. The authors are in debt to Li Rongcai, the chief-editor, Liang Jiaxin, the associate manager of Xidian Press for their great assistance in publication of the book.

目 录

第 1 章 绪论	1
1.1 人工神经网络概述	1
1.2 人工神经网络模型	4
1.2.1 神经元的数理模型	4
1.2.2 神经网络模型	7
1.3 二进前向神经网络	8
1.3.1 前向网络与二进前向网络	8
1.3.2 二进前向网络用于分类与分类超平面	9
1.3.3 训练样本与超立方体顶点着色、Boole 函数的关系	10
1.3.4 二进前向网络的性能指标	11
1.4 二进前向网络的研究现状	12
1.4.1 网络的分类能力	12
1.4.2 网络的分类行为与稳健性能	14
1.4.3 网络训练	14
1.5 本书的主要工作和特色	15
1.5.1 二进前向网络分类能力研究遇到的困难	15
1.5.2 本书的主要工作和内容安排	17
参考文献	18
 第 2 章 (稳健)感知器与(稳健)线性可分问题	21
2.1 感知器与线性可分问题	21
2.2 学习样本的变换	23
2.3 感知器学习算法	25
2.4 线性可分的充要条件	26
2.5 线性可分向量集的性质	28
2.6 稳健感知器的学习	29
2.7 非线性可分问题的结构	34
2.8 线性不可分问题学习的启发式算法	37
2.9 感知器学习算法学习线性不可分问题的基本性质	39
2.10 线性可分与非线性可分的对偶关系	40
参考文献	41
 第 3 章 二进前向网络的稳健/最稳健设计	44
3.1 引言	44
3.2 基本逻辑的稳健神经网络设计	45
3.2.1 逻辑与运算的稳健神经网络设计	46
3.2.2 逻辑或运算的稳健神经网络设计	46

3.2.3 逻辑非运算的稳健神经网络设计	47
3.3 网络结构与稳健性能	48
3.3.1 单隐层稳健网络设计及其稳健性能分析	49
3.3.2 Boole 函数的最稳健网络设计	50
3.4 仿真实验与结果	52
参考文献	55

第4章 稳健二进前向网络及其基本性质

4.1 引言	57
4.2 超立方体与超立方体图的构造	58
4.3 分类复杂度和稳健分类复杂度	60
4.4 稳健分类与稳健分类超平面	63
4.5 稳健分类超平面的辅阈值区域特性	65
4.6 稳健分类超平面辅阈值的反向贡献特性	68
4.7 稳健分类超平面方程的几种变形	69
4.8 稳健线性可分 Boole 函数的计数性质	70
4.9 基于稳健二进前向网络的逻辑规则提取	72
4.9.1 稳健神经元与 M of N 规则	73
4.9.2 稳健网络与规则网络	75
4.9.3 二进网络与规则网络	78
参考文献	79

第5章 稳健分类超平面系的基本模式及其(稳健)可分性

5.1 引言	82
5.2 两个稳健分类超平面的反向距、对偶原理	83
5.3 反向距与二阶稳健可分的关系	85
5.4 两个稳健分类超平面的辅阈值约束特性	87
5.5 分类超平面系的基本模式	88
5.6 分类超平面的反转和模式对调对隐空间分类超平面的影响	88
5.7 基本模式下隐空间分类超平面的构造	90
参考文献	94

第6章 分类超平面系的一般模式及其(稳健)可分性

6.1 引言	96
6.2 反向距概念的拓展	96
6.3 分类超平面系的一般模式	97
6.4 不相交模式分类超平面系的(稳健)可分性	99
6.5 相交模式分类超平面系的(稳健)可分性	102
6.6 几种典型分类超平面系的(稳健)可分性	105
参考文献	113

第7章 分类超平面系的退化和稳健退化

7.1 引言	114
--------------	-----

7.2 讨论退化条件的准备	115
7.2.1 退化的可能形式	116
7.2.2 红色顶点集之间的关系	117
7.2.3 使可能退化后的分类超平面的红色顶点集最小的条件	117
7.3 两个稳健分类超平面的退化	118
7.3.1 退化后的分类超平面形式	118
7.3.2 退化的充分条件	122
7.3.3 不可退化的必要条件	124
7.4 两个稳健分类超平面的稳健退化	126
7.4.1 稳健退化后的形式	126
7.4.2 稳健退化的充要条件	129
7.5 (稳健)分类超平面系的平行退化	132
7.5.1 (稳健)平行退化的引入	132
7.5.2 (稳健)分类超平面系的(稳健)平行退化条件	133
7.6 稳健分类超平面的构造	137
7.6.1 (稳健)网络学习对分类超平面系的要求	137
7.6.2 稳健分类超平面的构造条件	138
7.6.3 稳健分类超平面的构造规则	141
参考文献	146

第8章 二进前向网络的训练和学习 147

8.1 引言	147
8.2 多层感知器网络训练的反向传播算法	147
8.3 反向传播算法训练二进输入输出对的实验与结果	150
8.4 感知器网络的规则和非规则划分学习算法	154
8.5 二进前向网络的几何学习	157
8.5.1 线性可分问题的几何学习	157
8.5.2 非线性可分问题的几何学习算法	158
8.5.3 输出神经元的获得	159
8.6 稳健二进前向网络的几何学习	161
8.6.1 隐层稳健神经元的构造	162
8.6.2 平行模式下的稳健输出神经元	164
8.6.3 输出神经元的进化	166
8.6.4 实验与结果	168
参考文献	173

附录A 超立方体的有关理论 174

A.1 超立方体的概念与构造	174
A.2 超立方体中的平行线	177
A.3 子超立方体的数目	180
A.4 超立方体的谱	182
A.5 超立方体的特征分析	185
参考文献	187

Contents

Chapter 1 Introduction	1
1.1 Introduction to Artificial Neural Networks	1
1.2 Models of Artificial Neural Networks	4
1.2.1 Mathematical Models of Neurons	4
1.2.2 Models of Artificial Neural Networks	7
1.3 Binary Feedforward Neural Networks	8
1.3.1 Feedforward Neural Networks and Binary Feedforward Neural Networks	8
1.3.2 Binary Feedforward Neural Networks and Classification Hyperplanes	9
1.3.3 Training Samples, Coloring of Hypercube Vertices and Boolean Functions	10
1.3.4 Performance of Binary Feedforward Neural Networks	11
1.4 Outlines of Research Work on Binary Feedforward Neural Networks at Present	12
1.4.1 Classification Capabilities of Networks	12
1.4.2 Classification Behavior and Robust Performance	14
1.4.3 Network Training	14
1.5 Research Work and Characteristics of Present Book	15
1.5.1 Difficulties in Studying Classification Capability of Binary Feedforward Neural Networks	15
1.5.2 Main Works Presented in the Book and Their Arrangement	17
References	18
 Chapter 2 (Robust)Linearly Separable Problems and (Robust) Perceptrons	 21
2.1 Perceptron and Linearly Separable Problems	21
2.2 Transformation of Learning Samples	23
2.3 Perceptron Learning Algorithm	25
2.4 Necessary and Sufficient Condition for Linear Separability	26
2.5 Properties of Linearly Separable Training Sets	28
2.6 Robust Perceptron Learning Algorithm	29
2.7 Structure within Non-linearly Separable Training Sets	34
2.8 Heuristic Algorithm for Learning Maximal Number of Samples in a Linearly Inseparable Training Set	37
2.9 Analysis of Heuristic Algorithm for Linearly Inseparable Sets	39
2.10 Dual Relation of Linear and Non-linear Separability and Their Learning	40
References	41
 Chapter 3 Robust and Most Robust Design of Binary Feedforward Neural Networks	 44
3.1 Introduction	44

3.2 Robust Design of Neural Networks for Basic Logic	45
3.2.1 Robust Design of Neural Networks for AND	46
3.2.2 Robust Design of Neural Networks for OR	46
3.2.3 Robust Design of Neural Networks for NOT	47
3.3 Structure of designed Networks and Their Robust Performances	48
3.3.1 Robust Performance of Networks with One Hidden Layer	49
3.3.2 Most Robust Design of Networks for Realization of Boolean Functions	50
3.4 Simulations and Results	52
References	55
 Chapter 4 Robust Binary Feedforward Neural Networks and Their Properties	57
4.1 Introduction	57
4.2 Hypercube and Construction of Its Graph	58
4.3 Definitions of Classification Complexity and Robust Classification Complexity	60
4.4 Robust Classification and Robust Classification Hyperplanes	63
4.5 Region Characteristics of Complementary Threshold for a Robust Classification Hyperplane	65
4.6 Reverse Contribution of Complementary Threshold for a Robust Classification Hyperplane	68
4.7 Some Forms of Robust Classification Hyperplane Equation	69
4.8 Number of Robustly Linearly Separable Boolean Functions	70
4.9 Rule Extraction based on Robust Binary Feedforward Neural Networks	72
4.9.1 Robust Neurons and M of N Rules	73
4.9.2 Robust Networks and Rule Networks	75
4.9.3 Binary Networks and Rule Networks	78
References	79
 Chapter 5 Basic Formats of a System of Robust Classification Hyperplanes and Their (Robust) Separability	82
5.1 Introduction	82
5.2 Opposite Distance of a Pair of Robust Classification Hyperplanes and Dual Principle	83
5.3 Relation between Opposite Distance and Two-order Robust Separability	85
5.4 Complementary Threshold Constrains of a Pair of Robust Classification Hyperplanes	87
5.5 Basic Formats of a System of Classification Hyperplanes	88
5.6 Effect on Classification Hyperplane in Hidden Space with Classification Hyperplanes in Input Space Reversed and/or the Classes Exchanged	88
5.7 Construction of Classification Hyperplanes in Hidden Space with the System of Classification Hyperplanes in Input Space Being in Basic Formats	90
References	94

Chapter 6 General Formats of a System of Classification Hyperplanes and Their (Robust) Separability	96
6.1 Introduction	96
6.2 Extension of Opposite Distance	96
6.3 General Formats of a System of Classification Hyperplanes	97
6.4 (Robust) Separability of a System of Classification Hyperplanes under Non-intersect Format	99
6.5 (Robust) Separability of a System of Classification Hyperplanes under Intersect Format	102
6.6 Separability of Several Systems of Classification Hyperplanes	105
References	113

Chapter 7 Degeneration and Robust Degeneration of a System of Classification Hyperplanes	114
7.1 Introduction	114
7.2 Preparation for Studying Degeneration	115
7.2.1 Possible Forms of Degeneration	116
7.2.2 Relation Between Sets of Red Vertices	117
7.2.3 A Condition of Minimizing Sets of Red Vertices of Classification Hyperplanes after Possible Degeneration	117
7.3 Degeneration of a Pair of Robust Classification Hyperplanes	118
7.3.1 Form of Classification Hyperplane after Degeneration	118
7.3.2 Sufficient Condition for Degeneration	122
7.3.3 Necessary Condition for Non-degeneration	124
7.4 Robust Degeneration of a Pair of Robust Classification Hyperplanes	126
7.4.1 Form of Classification Hyperplanes after Robust Degeneration	126
7.4.2 Necessary and Sufficient Condition for Robust Degeneration	129
7.5 Parallel Degeneration of a System of (Robust) Classification Hyperplanes	132
7.5.1 Definition of Paralel Degeneration	132
7.5.2 (Robust)Parallel Degeneration Condition for a System of (Robust) Classification Hyperplanes	133
7.6 Construction of Robust Classification Hyperplanes	137
7.6.1 Requirements of Robust Learning to a System of Classification Hyperplanes	137
7.6.2 Construction Conditions of Robust Classification Hyperplanes	138
7.6.3 Construction Rules of Robust Classification Hyperplanes	141
References	146

Chapter 8 Training and Learning of Binary Feedforward Neural Networks	147
8.1 Introduction	147
8.2 Back-propagation Algorithm for Training Multilayer Feedforward Neural Networks	147
8.3 Experiments and Results on Training Networks with Back-propagation Algorithm for Binary Input-Output Training Samples	150
8.4 Regular and Irregular Partitioning Algorithms for Learning of Perceptron Networks	154

8.5	Geometric Learning of Binary Feedforward Neural Networks	157
8.5.1	Geometric Learning of Linearly Separable Problems	157
8.5.2	Geometric Learning of Non-linearly Separable Problems	158
8.5.3	Output Neuron Obtaining	159
8.6	Geometric Learning of Robust Binary Feedforward Neural Networks	161
8.6.1	Construction of Robust Hidden Neurons	162
8.6.2	Finding Robust Output Neuron in the Case of Parallel Format of Robust Classification Hyperplanes in Input Space	164
8.6.3	Evolution of Robust Output Neuron	166
8.6.4	Experiments and Results	168
	References	173
 Appendix A Theory of Hypercubes		174
A.1	Hypercubes and Construction of Their Graphs	174
A.2	Parallel Lines in Hypercubes	177
A.3	Number of Sub-hypercubes	180
A.4	Spectrum of Hypercubes	182
A.5	Eigenvector Analysis of Hypercube Graphs	185
	References	187

第 1 章

绪 论

本章简要概述了人工神经网络，介绍了人工神经网络模型的不同形式，并着重对二进前向神经网络的特点、研究现状和存在问题作了简短回顾。

1.1 人工神经网络概述

大家知道，人是地球上具有最高智慧的生物，而人的智能均来自大脑，人类靠大脑进行思考、联想、推理和判断，这些功能是何被称为“电脑”的一般计算机所无法取代的。长期以来，很多科学家一直致力于人脑内部结构和功能的探讨研究，并试图建立模仿人类大脑的计算机，直到目前人们对大脑的内部工作机理还不甚清楚，但对其结构已经有所了解。粗略地讲，大脑是由大量神经细胞或神经元组成的。每个神经元可看作一个小的处理单元，这些神经元按某种方式互相连接起来，形成大脑内部的生理神经网络，这些神经网络中各神经元之间连接的强弱随外部的激励信号作自适应变化，而每个神经元又随着所接收到多个激励信号的综合大小而呈现兴奋或抑制状态。现已明确大脑的学习过程就是神经元之间连接强度随外部激励信息作自适应变化的过程，而大脑处理信息的结果则由神经元的状态表现出来。显然我们建立的信息处理系统实际就是模仿生理神经网络，所以称它为人工神经网络。同时也必须注意，尽管人工神经网络是大脑结构的模仿，但这种模仿目前还处于极低的水平。为了更确切起见，美国神经网络学家 Hecht Nielsen 给出了如下定义：神经网络是由多个非常简单的处理单元彼此按某种方式相互连接而形成的计算系统，该系统是靠其状态对外部输入信息的动态响应来处理信息的。

人工神经网络的操作有两种过程：一是学习或训练，二是正常操作或称为联想。训练时，把要教给网络的信息(外部输入)作为网络的输入和要求的输出，使网络按某种规则(称为训练算法或学习

算法)调节各神经元之间的连接权数值,直到加上给定输入网络,就能产生给定输出为止。这时,各个连接权已经调节好,网络的训练过程完成。所谓正常操作,就是对训练好的网络输入一个信号,它就可正确联想出相应的输出,所以也称为联想操作。就像小孩认人一样。

尽管人工神经网络还是生物大脑低水平的模仿,却在图像识别、语音识别、联想记忆、预测和优化等方面表现了很好的智能特性和极好的应用前景。

人工神经网络是基于人类大脑的结构和功能而建立起来的新学科。尽管目前它只是大脑的低级模仿,但它的很多特点和人类的智能特点类似。

(1) 它不但结构上是并行的,它的处理顺序也是并行的和同时的。在同一层内的处理单元都是同时操作的,即神经网络的计算功能分布在各处理单元上。目前的神经网络常常用一般计算机的串行工作方式模拟它的并行处理方式,所以显得很慢,而真正的神经网络将会大大提高处理速度,并能实现实时处理。

(2) 知识的分布存储。在神经网络中,知识不是存储在特定的存储单元中,而是分布在整个系统中,要存储多个知识就需要很多连接。在神经网络中要获得存储的知识则需采用“联想”的办法,这类似人类和动物的联想记忆。当给一个神经网络输入一个激励时,它要在已存储的知识中寻找与该输入匹配最好的存储知识作为其解,这犹如教师可以容易地辨认出他的学生的笔迹,即使这些笔迹是变形的、失真的。这就是说,人类根据联想善于正确地进行识别,人工神经网络也是这样。

(3) 容错性和稳健性。人类大脑具有很强的容错能力。我们知道,每天大脑的一些细胞会自动死去,但并没有影响人们记忆和考虑问题的能力,同时,人类大脑还具有非常强的对数据处理的稳健能力,可以从不完善和不完整的数据和图像中进行学习并作出决定。这正是大脑中知识是存储在很多处理单元和它们的连接上的结果。人工神经网络也是如此,由于知识存在于整个系统中,而不是仅仅存在于一个存储单元中,一定比例的单元不参与运算,对整个系统的性能不会产生重大影响,数据的不完整性和不完善性也不会对整个系统的性能产生重大影响。

(4) 自适应性。人类有很强的适应外部环境的能力。孩子在周围环境的熏陶下可以学会认字、说话、走路以及考虑判断等,人工神经网络也可以通过学习具备这种能力。网络这种能力的获得可以通过两种形式:一是有指导的训练,一是无指导的训练。对网络的有指导训练是通过“教师”对网络的指导“教会”网络做如分类、识别等工作。对网络的无指导训练是通过网络自身的自组织,即通过网络的连接权依照某一准则的自适应调整,实现对输入数据的分类、识别等工作。网络的学习训练是一个艰难的过程。目前网络的设计没有严格确定的方法(一般凭经验确定),所以选择训练方法和所需网络结构没有统一标准。脱机训练往往需要很长的时间,为了获得最佳效果,常常需要重复实验多次才行。

人工神经网络能在一定程度上模仿生物神经系统的智慧和功能,使得它在信息处理和智能科学中有着极其重要的地位,世界各国的科学家们对神经网络这门科学给予了极大的关注,并且成为国际前沿研究的热点。例如:美国国防部预研局(Defense Advanced Research Projects Agency)资助众多科学家和军事专家,经过多方严格而周密的论证,决定把基于神经网络的自动目标识别、声纳信号处理、语音识别、地震勘探信号处理等作为研究的重点,并认为它们是最有前途且能取得重大突破的应用领域;日本政府也积极支持这